

**SEMINAR NASIONAL PENDIDIKAN SAINS**

“Strategi Pengembangan Pembelajaran dan Penelitian Sains untuk Mengasah Keterampilan Abad 21
Creativity and Innovation, Critical Thinking and Problem Solving, Communication, Collaboration/4C)”
Universitas Sebelas Maret Surakarta, 26 Oktober 2017



ANALISIS VALIDITAS ISI *COMPUTERIZED TWO-TIER MULTIPLE CHOICE* UNTUK MENGUKUR LITERASI SAINS SISWA

Marantika Lia Kristiyasari¹, Sri Yamtinah², Suryadi Budi Utomo³

¹ Magister Pendidikan Sains FKIP Universitas Sebelas Maret, Surakarta, 57126

² Magister Pendidikan Sains FKIP Universitas Sebelas Maret, Surakarta, 57126

³ Magister Pendidikan Sains FKIP Universitas Sebelas Maret, Surakarta, 57126

Email korespondensi: christmarli417@gmail.com

Abstrak

Penelitian ini bertujuan untuk menganalisis validitas isi instrumen penilaian *Computerized Two-Tier Multiple Choice* (CTTMC) untuk mengukur literasi sains siswa pada materi IPA Terpadu SMP. Instrumen CTTMC merupakan instrumen penilaian berupa soal pilihan ganda bertingkat dua yang terdiri dari *first tier* (berisi pilihan jawaban) dan *second tier* (alasan pemilihan jawaban). Analisis validitas isi instrumen penilaian ini menggunakan formula Aiken (1985). Metode yang digunakan dalam penelitian ini adalah metode deskriptif kuantitatif, berdasarkan hasil perhitungan validitas isi dari formula Aiken. Data validitas isi diperoleh dari 9 validator yang terdiri 3 dosen ahli (bidang biologi, fisika, kimia, dan evaluasi), sedangkan 6 orang lainnya berasal dari guru IPA SMP di kota Surakarta melalui *forum group discussion* (FGD). Berdasarkan hasil penelitian menunjukkan bahwa 1) nilai validitas isi tertinggi adalah 1, 2) indeks validitas isi (V_{tabel}) diperoleh 0.74, 3) prosentase validitas isi sebesar 80%, 4) diperoleh 24 item soal CTTMC untuk mengukur literasi sains yang dinyatakan valid dari 30 item soal yang di validasi.

Kata kunci: FGD, formula Aiken, IPA Terpadu, instrumen penilaian

Pendahuluan

Pendidikan merupakan salah satu sektor yang penting dalam pembangunan di setiap Negara. Menurut Undang-Undang Nomor 20 Tahun 2003 pasal 1, Pendidikan adalah usaha sadar dan terencana mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara. Pendidikan nasional berfungsi mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang bermartabat dalam rangka mencerdaskan kehidupan bangsa, bertujuan untuk berkembangnya potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu cakap, kreatif, mandiri dan menjadi warga negara yang demokratis serta bertanggung jawab (UU RI No. 20 Tahun 2003 Pasal 3).

Perkembangan ilmu pengetahuan dan teknologi telah membawa perubahan yang besar di dalam kehidupan manusia dan menimbulkan adanya persaingan secara global. Tantangan dan perkembangan pendidikan di Indonesia pada masa sekarang maupun nanti akan menjadi semakin kompleks. Hal ini disebabkan oleh adanya perubahan tuntutan masyarakat terhadap kualitas pendidikan itu sendiri. Dengan demikian, pendidikan diharapkan dapat menjadi jembatan yang dapat menghubungkan individu dengan lingkungannya di era globalisasi yang semakin berkembang dengan pesat sekarang ini. Sehingga individu tersebut mampu berperan sebagai SDM yang berkualitas.

Tiga bagian yang sangat penting di dalam pendidikan adalah kurikulum, proses pembelajaran, penilaian. Kurikulum merupakan penjabaran dari tujuan pendidikan yang menjadi acuan dalam proses pembelajaran. Proses pembelajaran merupakan suatu proses yang bertujuan agar peserta didik dapat belajar melalui perencanaan dan pengaturan lingkungan, sarana, dan prasarana yang mendukung terwujudnya kegiatan pembelajaran. Penilaian merupakan kegiatan yang dilakukan untuk mengukur dan menilai tingkat ketercapaian dari kurikulum yang digunakan.

Salah satu bentuk penilaian dalam kegiatan pembelajaran adalah ulangan harian. Melalui kurikulum 2013 saat ini, diharapkan fungsi dan tujuan pendidikan nasional dapat tercapai. Kurikulum 2013 memfasilitasi dan mendorong guru maupun peserta didik untuk mendapatkan Standar Kompetensi Lulusan yang diinginkan. Setiap guru dituntut untuk mengaplikasikan kegiatan pembelajaran dengan menggunakan pendekatan saintifik. Pendekatan saintifik ini cocok untuk seluruh mata pelajaran, terutama untuk mata pelajaran yang berhubungan dengan ilmu pengetahuan alam atau sains. Dalam kurikulum 2013 ini, penilaian siswa dalam proses pembelajaran sangat erat kaitannya dengan keterampilan berpikir. Keterampilan berpikir siswa dalam memahami, menganalisis dan mendekonstruksi informasi pada pembelajaran sains dapat dilatih melalui pengembangan kemampuan literasi sains.

Literasi sains (*Scientific literacy*) merupakan salah satu ranah studi *Programme for International Student Assessment* (PISA) yang terfokus pada kemampuan seseorang untuk dapat menggunakan pengetahuan dan keterampilan sains dalam membuat keputusan dan memecahkan suatu permasalahan. Literasi sains merupakan pengetahuan ilmiah individu dan kapasitas menggunakan pengetahuan tersebut untuk mengidentifikasi pertanyaan, memperoleh pengetahuan baru, menjelaskan fenomena ilmiah dan menarik kesimpulan berdasarkan bukti tentang isu-isu yang berkaitan dengan sains (OECD, 2014).

National Teacher Association (1998) mengemukakan bahwa seorang yang literat sains adalah orang yang menggunakan konsep sains, keterampilan proses dan nilai dalam membuat keputusan sehari-hari yang berhubungan dengan orang lain maupun lingkungan dan memahami interelasi antara sains, teknologi dan masyarakat termasuk dalam hal perkembangan sosial dan ekonomi.

Pembelajaran sains sangat bertanggung jawab atas literasi sains peserta didik, sehingga kualitas pendidikan sains perlu ditingkatkan agar dapat mencapai taraf pengembangan yang berkelanjutan (Liliasari, 2011).

Tiga elemen standar yang terdapat dalam literasi sains, antara lain proses sains, konten sains, dan konteks aplikasi sains. Kemampuan literasi sains setiap individu penting untuk dikembangkan dan dikuasai oleh peserta didik, terutama dalam memahami lingkungan hidup, memiliki kemampuan dan kreativitas individu untuk menggunakan pengetahuan serta keterampilan ilmiahnya dalam memecahkan masalah terutama yang berhubungan dengan kehidupan sehari-hari, serta dapat membuat keputusan yang dapat dipertanggungjawabkan. Tujuannya agar setiap individu mampu menghadapi tantangan global yang semakin kompleks.

Instrumen *Computerized Two-tier Multiple Choice* (CTTMC) merupakan instrumen *Two-tier Multiple Choice* (TTMC) yang berbasis komputer (terkomputerisasi). Instrumen *Two-tier Multiple Choice* (TTMC) merupakan suatu instrumen penilaian berupa tes objektif dua tingkat. Tingkatan pertama berupa soal utama (*first tier*) dan tingkatan kedua berupa alasan pemilihan jawaban (*second tier*). Ditunjang dengan model penskoran *Graded Response Model* (GRM), maka guru akan dengan mudah melakukan pemeriksaan jawaban siswa namun sekaligus dapat mendeteksi kemampuan siswa (Wardani, *et al.*, 2015). Tuysuz (2009) menunjukkan bahwa tes pilihan ganda dua tingkat ini sangat efektif untuk menentukan miskonsepsi peserta didik dan dapat digunakan sebagai alternatif penilaian tes evaluasi terhadap prestasi belajar peserta didik. Embertson dan Raise (2000: 97) menyatakan bahwa *Graded Response Model* (GRM) sesuai digunakan ketika item yang dipakai dapat dikategorikan atau memiliki respon bertingkat seperti yang digunakan pada skala Likert.

Penerapan *Graded Response Model* (GRM) ini diaplikasikan pada penskoran instrumen *Computerized Two-Tier Multiple Choice* (CTTMC). Aplikasi dari GRM pada instrumen tersebut tersaji pada tabel 1.

Tabel 1: Skoring Pada Instrumen <i>Computerized Two-Tier Multiple Choice</i> (CTTMC)		
No	Aspek Penilaian	Skor
1	Jawaban benar pada <i>First tier</i> dan <i>Second tier</i>	3
2	Jawaban benar pada <i>First tier</i> dan jawaban salah pada <i>Second tier</i>	2
3	Jawaban salah pada <i>First tier</i> dan jawaban benar pada <i>Second tier</i>	1
4	Jawaban salah pada <i>First tier</i> dan <i>Second tier</i>	0

Pedoman penskoran tersebut digunakan untuk menganalisis kemampuan literasi sains siswa. Selain itu, dengan menggunakan GRM dapat pula diketahui kelemahan siswa, sehingga akan dapat diketahui profil kemampuan setiap siswa berdasarkan jawaban yang mereka berikan.

Ada beberapa pendapat mengenai validitas untuk instrumen yang digunakan pengukuran, baik di dalam bidang pendidikan maupun psikologi. Menurut *American Educational Research Association* (AERA), *American Psychological Association* (APA), and *National Council on Measurement in Education* (NCME) dalam *Standards for Educational and Psychological Testing* (1999), menyatakan bahwa validitas merujuk pada derajat dari fakta dan teori yang mendukung interpretasi skor tes, serta merupakan pertimbangan paling penting dalam pengembangan tes. Ahli lain mengemukakan bahwa validitas suatu alat ukur adalah sejauh mana alat ukur itu mampu mengukur apa yang seharusnya diukur (Nunnally, 1978; Allen & Yen, 1979: 97; Kerlinger, 1986).

Validitas dapat dikelompokkan menjadi tiga tipe, yaitu: (1) validitas kriteria/*criterion-related validity*, (2) validitas isi/*content validity*, dan (3) validitas konstruk/*construct validity* (Nunnally, 1978; Allen & Yen, 1979: 97; Fernandes, 1984; Kerlinger, 1986; dan Lawrence 1994).

Bollen (1989) mendefinisikan validitas isi sebagai sebuah validasi kualitatif dimana domain konsep menjadi lebih jelas dan tugas ahli adalah menganalisis apakah sudah secara keseluruhan mewakili domain yang diukur. Dengan demikian, validitas isi merupakan sarana kualitatif untuk memastikan bahwa indikator telah sesuai dengan konsep yang di definisikan oleh peneliti.

Validitas isi (*content validity*) suatu instrumen adalah sejauh mana butir-butir dalam instrumen itu mewakili komponen-komponen dalam keseluruhan kawasan isi dari objek yang hendak diukur tersebut dan sejauh mana butir-butir itu mencerminkan ciri perilaku yang hendak diukur (Nunnally, 1978; Fernandes, 1984). Sementara itu Lawrence (1994) menjelaskan bahwa validitas isi merupakan keterwakilan pertanyaan terhadap kemampuan khusus yang harus diukur. Berdasarkan penjelasan di atas, maka dapat disimpulkan bahwa validitas isi terkait dengan analisis rasional kemampuan yang hendak diukur. Sehingga, dengan demikian validitas isi menjadi hal yang penting untuk dianalisis dalam instrumen penilaian, terutama instrumen penilaian *Computerized Two-Tier Multiple Choice* (CTTMC) untuk mengukur kemampuan literasi sains.

Metode Penelitian

Penelitian ini menggunakan metode deskriptif kuantitatif. Desain penelitian deskriptif kuantitatif digunakan dengan tujuan untuk menjelaskan suatu situasi yang hendak diteliti dengan dukungan studi kepustakaan sehingga lebih memperkuat analisa peneliti dalam suatu kesimpulan. Hasil penelitian yang diperoleh berasal dari hasil perhitungan indikator-indikator variable penelitian yang kemudian dipaparkan secara tertulis.

Hasil Penelitian dan Pembahasan

Validitas isi dapat ditentukan dengan melalui dua tahapan: pengembangan dan *judgment* ahli (Lynn, 1986). Dalam pemilihan ahli atau validator harus sesuai dengan instrumen yang akan ditelaah. Pemilihan ahli ini sangat penting karena akan menentukan kualitas dari instrumen yang dikembangkan.

Pada penelitian ini roses validitas ini dilakukan pada *Forum Group Discussion* (FGD) yang melibatkan 9 ahli (*expert/validator*). Validator yang dipilih adalah 3 dosen ahli (bidang biologi, fisika, kimia dan instrumen) yang telah bergelar doktor. Sedangkan 6 orang lainnya berasal dari guru IPA SMP yang telah berpengalaman mengajar lebih dari 15 tahun dan bergelar magister.

Lembar validasi terdiri dari 3 aspek utama, yaitu aspek materi, konstruksi, dan aspek bahasa/budaya yang dijabarkan menjadi 14 item. Para ahli (*expert*) diminta untuk melakukan telaah pada setiap soal CTTMC untuk mengukur kemampuan literasi sains siswa. Hasil validasi dari 9 orang ahli (*expert*) ini akan dianalisis menggunakan formula Aiken. Sehingga diperoleh nilai validitas isi dari instrumen penilaian CTTMC.

Terdapat empat pilihan kategori dalam lembar validasi instrumen. Berdasarkan penelitian Shidiq (2016), instrumen dikatakan relevan (R) apabila dari sisi konstruksi, materi dan bahasa sudah


baik. Instrumen dikatakan cukup relevan (CR) apabila indikator, keterampilan, kemampuan dan materi yang akan diukur sudah sesuai namun masih terdapat kesalahan dalam pemilihan kata (bahasa). Instrumen dikatakan kurang relevan (KR) apabila kemampuan dan keterampilan atau materi yang hendak diukur kurang cocok (perlu direvisi), dan instrumen dikatakan tidak relevan (TR) apabila butir soal sama sekali tidak mencerminkan indikator soal, indikator kemampuan literasi sains, dan kompetensi dasar yang akan diukur.

Para ahli diminta untuk memberikan tanda cek (✓) di setiap butir soal yang telah ditelaah pada tabel yang telah disusun oleh peneliti seperti yang terlihat pada Gambar 1.

**VALIDASI INSTRUMEN PENILAIAN “COMPUTERIZED TWO TIER MULTIPLE CHOICE”
UNTUK MENGUKUR LITERASI SAINS**

KD: 3.1. Menganalisis gerak pada makhluk hidup, sistem gerak pada manusia, dan upaya menjaga kesehatan sistem gerak

KD: 3.3. Menjelaskan konsep usaha, pesawat sederhana, dan penerapannya dalam kehidupan sehari-hari termasuk kerja otot pada struktur rangka manusia

No	Indikator Literasi Sains	Indikator Soal	No Soal	Butir Soal (First Tier) dan Pilihan Alasan (Second Tier)	Validasi				Saran
					TR	KR	CR	R	
1.	Mengidentifikasi pertanyaan ilmiah	Disajikan sebuah aktivitas yang dilakukan oleh beberapa orang, ada yang sedang menimba air, memotong kemiri, membuka tutup botol, memasukkan drum ke dalam bak truk, dan mencabut paku. ❖ Peserta didik dapat menentukan urutan jenis-jenis dari pesawat sederhana ❖ Peserta didik dapat menyebutkan jenis-jenis pesawat sederhana secara urut	15	Perhatikan gambar di bawah ini! Aktivitas ini menggunakan prinsip kinerja otot dan merupakan jenis pesawat sederhana dalam kehidupan sehari-hari.  Jenis-jenis pesawat sederhana secara urut, ditunjukkan pada nomor a. 5,2,3,4,1 b. 3,5,2,4,1 c. 2,3,5,4,1 d. 5,3,2,4,1 Karena, nomor tersebut merupakan: a. (3) Tuas golongan pertama, (5) tuas golongan kedua, (2) tuas golongan ketiga, (4) bidang miring dan (1) katrol b. (2) Tuas golongan pertama, (3) tuas golongan kedua, (5) tuas golongan ketiga, (4) bidang miring dan (1) katrol c. (5) Tuas golongan pertama, (3) tuas golongan kedua, (2) tuas golongan ketiga, (4) bidang miring dan (1)					

Gambar 1. Tabel Validasi Instrumen *Computerized Two-Tier Multiple Choice* (CTTMC) untuk mengukur kemampuan literasi sains

Pada Gambar 1, tabel validasi soal terdiri dari indikator soal, indikator kemampuan literasi sains, kategori validasi serta saran. Pada kolom kategori, ahli hanya memberikan tanda cek (✓) pada kategori yang sesuai. Pada kolom saran, ahli dapat memberikan saran secara umum maupun terkait dengan item soal yang sedang ditelaah. Melalui bentuk tabulasi pada Gambar 1, ahli dapat melihat keterkaitan antara indikator soal, soal dan indikator kemampuan literasi sains. Total butir soal yang divalidasi oleh ahli adalah 30 soal.

Indikator kemampuan literasi sains yang ditelaah dalam instrumen *Computerized Two-Tier Multiple Choice* (CTTMC) ini meliputi mengidentifikasi pertanyaan ilmiah, mengidentifikasi bukti yang diperlukan dalam penyelidikan ilmiah, menarik dan mengevaluasi kesimpulan, mengkomunikasikan kesimpulan yang valid, dan menggunakan konsep-konsep sains dalam situasi yang berbeda dari apa yang dipelajarinya.

Pada kategori validasi, terdapat 4 kategori yaitu relevan (R) yang bernilai 4, cukup relevan (CR) bernilai 3, kurang relevan (KR) bernilai 2, dan tidak relevan (TR) bernilai 1. Hasil analisis validasi setiap soal kemudian dihitung menggunakan formula Aiken (1985) seperti beriku ini rumusnya:

$$V = \frac{S}{[n \times (c-1)]} ; S = \sum n_i(r-l_o)$$

V menyatakan indeks validitas dari Aiken, c adalah banyaknya kategori atau kriteria, l_o merupakan kategori terendah (dalam hal ini kategori tidak relevan (TR) yang termasuk kategori terendah karena bernilai 1), n_i adalah banyaknya penilai (*raters*) yang memilih kriteria i , dan r merupakan kriteria ke i , serta n menyatakan jumlah keseluruhan penilai (*raters*).

Nilai V berkisar 0-1 dan kriteria yang digunakan untuk menyatakan sebuah butir soal dikatakan valid secara isi pada jumlah rater (penilai) sebanyak 9 orang berdasarkan tabel Aiken adalah 0,74 (Aiken, 1985). Setiap item soal dianalisis dengan menggunakan cara yang sama. Hasil analisis setiap soal kemudian ditabulasikan ke dalam tabel analisis kualitatif seperti tampak pada Gambar 2. Jika pada Gambar 1, setiap item soal dianalisis berdasarkan kategori validasi dan keterkaitan dengan indikator soal, soal dan indikator kemampuan literasi sains. Pada Gambar 2 ini, item soal dikelompokkan sesuai aspeknya. Ahli memberikan tanda cek (✓) pada aspek yang sesuai dan tanda silang (x) pada aspek yang tidak sesuai.

Hasil analisis ahli kemudian dianalisis secara kualitatif dan dikelompokkan menjadi 3 aspek yaitu aspek materi, aspek konstruksi, dan aspek bahasa/budaya seperti yang tersaji pada Gambar 2.

TELAH KUALITATIF INSTRUMEN PENILAIAN "COMPUTERIZED TWO TIER MULTIPLE CHOICE" UNTUK MENGUKUR LITERASI SAINS																																
No	Aspek yang ditelaah	Nomor Soal																														
		1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	
A. Materi																																
1.	Soal sesuai dengan indikator																															
2.	Materi yang dinyatakan sesuai dengan kompetensi																															
3.	Pilihan jawaban homogen dan logis																															
4.	Hanya ada satu kunci jawaban																															
B. Konstruksi																																
5.	Pokok soal dirumuskan dengan singkat, jelas dan tegas																															
6.	Rumusan pokok soal																															
	dan pilihan jawaban merupakan pernyataan yang diperlukan saja																															
7.	Pokok soal tidak memberi petunjuk kunci																															
8.	Pokok soal bebas dari pernyataan yang bersifat negatif																															
9.	Pilihan jawaban yang berbentuk angka atau waktu harus disusun berdasarkan urutan besar kecilnya nilai angka tersebut atau kronologis waktunya																															
10.	Gambar, grafik, tabel, diagram dan sejenisnya yang terdapat pada soal harus jelas dan berfungsi																															
C. Bahasa/Budaya																																
11.	Menggunakan bahasa yang sesuai kaidah bahasa Indonesia																															
12.	Menggunakan bahasa yang komunikatif																															
13.	Tidak menggunakan bahasa yang berakur																															
	terempat tabu																															
14.	Pilihan jawaban tidak mengulang kata kelompok kata yang sama, kecuali merupakan satu kesatuan pengertian																															

Gambar 2. Tabel telaah kualitatif instrumen penilaian *Computerized Two-Tier Multiple Choice* (CTTMC)

Lembar validasi yang telah diisi oleh 9 orang ahli (*expert*) kemudian ditabulasikan dan dihitung menggunakan formula Aiken dengan bantuan program *Microsoft Excel*. Setiap item soal mendapatkan skor validasi dari 9 ahli. Hasil perhitungan validasi dengan menggunakan validasi Aiken dapat dilihat pada Tabel 2.

Tabel 2. Hasil perhitungan validitas isi menggunakan formula Aiken

No Soal	V	V _{tabel}	Kesimpulan
1	0.70	0.74	TIDAK VALID
2	0.48	0.74	TIDAK VALID
3	0.96	0.74	VALID
4	1	0.74	VALID
5	1	0.74	VALID
6	0.93	0.74	VALID
7	0.81	0.74	VALID
8	0.85	0.74	VALID
9	0.81	0.74	VALID
10	0.67	0.74	TIDAK VALID
11	0.89	0.74	VALID
12	0.93	0.74	VALID
13	0.85	0.74	VALID
14	0.93	0.74	VALID
15	0.89	0.74	VALID
16	0.85	0.74	VALID
17	0.78	0.74	VALID
18	0.85	0.74	VALID
19	0.85	0.74	VALID
20	0.63	0.74	TIDAK VALID
21	0.74	0.74	VALID
22	0.67	0.74	TIDAK VALID
23	0.93	0.74	VALID
24	0.70	0.74	TIDAK VALID
25	0.89	0.74	VALID
26	0.89	0.74	VALID
27	0.89	0.74	VALID
28	0.96	0.74	VALID
29	0.74	0.74	VALID
30	0.96	0.74	VALID

Berdasarkan perhitungan nilai V di atas pada Tabel 2, diperoleh hasil antara lain: (1) V_{tabel} dari validitas Aiken sebesar 0.74. Ini dilihat dan dibaca dari tabel validitas Aiken, dengan jumlah penilai atau *raters* 9 orang dan terdiri dari 4 kategori penilaian validasi. (2) nilai V berdasarkan perhitungan dengan menggunakan Ms.Excel, V maksimum sebesar 1. (3) dari 30 soal yang di validasi oleh *expert*, 24 soal yang dinyatakan valid dan 4 soal dinyatakan tidak valid, sehingga soal yang dinyatakan tidak valid perlu dilakukan revisi. Hal ini dikarenakan 4 soal yang dinyatakan tidak valid, nilai perhitungan validasinya kurang dari 0.74 (V_{tabel}). (4) prosentase validitas isi sebesar 80%.

Instrumen yang valid menjadi bagian penting dalam proses penilaian karena instrumen yang valid menjadi tujuan yang fundamental bagi pengembangan instrumen. Instrumen yang valid mampu mengukur apa yang seharusnya diukur dalam hal ini mengukur kemampuan kognitif siswa dan kemampuan literasi sains siswa.

Mengembangkan instrumen yang reliabel dan valid merupakan proses yang panjang. Namun, yang paling dasar instrumen harus valid secara isi. Hal ini karena item soal yang dihasilkan akan mewakili konstruk yang akan diukur. Oleh sebab itu, validitas isi menjadi kunci utama suatu instrumen yang berkualitas.

Simpulan, Saran, dan Rekomendasi

Berdasarkan hasil penelitian dan pembahasan dapat disimpulkan bahwa terdapat 24 item soal *Computerized Two-Tier Multiple Choice* (CTTMC) untuk mengukur kemampuan literasi sains siswa yang dinyatakan valid berdasarkan validitas Aiken, sedangkan ada 4 item soal yang dinyatakan tidak valid, sehingga perlu adanya revisi. Soal CTTMC yang disusun telah mewakili indikator literasi sains yang akan diukur. Dengan demikian, prosentase keseluruhan validitas isi pada instrumen soal CTTMC sebesar 80%.

Dari hasil penelitian ini, diharapkan mampu menjadi pengetahuan baru bagi pembaca. Selain itu, diharapkan juga bisa menjadi referensi bagi penelitian mengenai analisis validitas isi instrumen penilaian atau penelitian-penelitian lainnya yang sejenis.

Daftar Pustaka

- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45, 131-142
- Allen, M. J. & Yen, W. M. (1979). *Introduction to measurement theory*. Monterey, CA: Brooks/Cole Publishing Company
- American Educational Research Association, American Psychological Association, and National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Psychological Association
- Bollen, K. A. (1989). *Structural Equation with Latent Variables* (pp. 179-225). John Wiley & Sons.
- Departemen Pendidikan Nasional. (2003). *Undang-Undang No 23 Tahun 2003, Tentang Sistem Pendidikan Nasional*. Jakarta: Depdiknas
- DeVellis, R. F. (2003). “Scale development: Theory and Application,” (2nd Ed.), United Kingdom: Sage
- Fernandes, H. J.X. (1984). *Evaluation of educational program*. Jakarta: National Education Planning, Evaluating and Curriculum Development
- Harlen, W. (2013). *Assesment & Inquiry-Based Science Education: Issues in Policy and Practice*
- Kerlinger, F. N. (1986). *Asas-asas penelitian behavioral* (Terjemahan L.R. Simatupang). Yogyakarta: Gajahmada University Press
- Lawrence, M. R. (1994). Question to ask when evaluating test. *Eric Digest. Artikel*. Diambil dari: <http://www.ericfacility.net/ericdigest/ed.385607.html>
- Liliasari. (2011). *Membangun Masyarakat Melek Sains Berkarakter Bangsa Melalui Pembelajaran. Makalah Seminar Nasional UNNES*. [Online]. Tersedia: <http://liliasari.staf.upi.edu/files/2011/05/Makalah-Semnas-UNNES-2011.Liliasari.pdf>. [7 Januari 2017]
- Lynn, M. (1986). Determination and quantification of content validity. *Nursing Research*, 1986, 35 (6), 382-385
- National Science Teachers Association (NSTA). (1998). NSTA Standards for Science Teachers Preparation Adopted by the NSTA Board of Directors, 1998. Tersedia: <http://www.nvgc.vt.edu/nsta-ncate/november98.htm>.
- Nunnally, J. (1978). *Psychometric theory* (2nd ed.). New York: Mc Graw Hill

- OECD. (2006). *Assesing Scientific, Reading and Mathematic Literacy A Framewowk for PISA 2006*. Paris: *OECD Publications*
- Programme for International Student Assessment (PISA). (2012). *PISA 2012 Plus Results Performance of 15-years-olds in reading, mathematics and science for 10 additional participants*. [Online]. Tersedia: <http://nces.ed.gov/surveys/pisa>
- Tuyuz, C. (2009). Development of Two-Tier Diagnostic Instrument and Assess Students' Understanding in Chemistry. *Scientific Research and Essay* Vol. 4 (6) pp. 626-631. ISSN 1992-2248